

Error Covariance Calculation for Forecast Bias Estimation in Hydrologic Data Assimilation

Valentijn R. N. Pauwels

Monash University, Department of Civil Engineering, Clayton, Victoria, Australia

Gabriëlle J. M. De Lannoy

NASA Goddard Space Flight Center, Greenbelt, Maryland

Abstract

To date, an outstanding issue in hydrologic data assimilation is a proper way of dealing with forecast bias. A frequently used method to bypass this problem is to rescale the observations to the model climatology. While this approach improves the variability in the modeled soil wetness and discharge, it is not designed to correct the results for any bias. Alternatively, attempts have been made towards incorporating dynamic bias estimates into the assimilation algorithm. Persistent bias models are most often used to propagate the bias estimate, where the a priori forecast bias error covariance is calculated as a constant fraction of the unbiased a priori state error covariance. The latter approach is a simplification to the explicit propagation of the bias error covariance. The objective of this paper is to examine to which extent the choice for the propagation of the bias estimate and its error covariance influence the filter performance. An Observation System Simulation Experiment (OSSE) has been performed, in which ground water storage observations are assimilated into a biased conceptual hydrologic model. The magnitudes of the forecast bias and state error covariances are calibrated by optimizing the innovation statistics of groundwater storage. The obtained bias propagation models are found to be identical to persistent bias models. After calibration, both approaches for the estimation of the forecast bias error covariance lead to similar results, with a realistic attribution of error variances to the bias and state estimate, and significant reductions of the bias in both the estimates of groundwater storage and discharge. Overall, the results in this paper justify the use of the traditional approach for online bias estimation with a persistent bias model and a simplified forecast bias

error covariance estimation.

Keywords: Data Assimilation, Bias, Kalman Filter

1. Introduction

Hydrologic models, which simulate the partitioning of rainfall into infiltration and surface runoff, and the partitioning of the net radiation into sensible, latent and ground heat fluxes, have been used for decades. Due to errors or uncertainties in the meteorological forcing data, the parameters, the land cover, soil, and topographic input data, and the model formulation, these models are prone to errors. One way of correcting the model results is to periodically update the model using externally obtained data (i.e. independent from the modeling system), which is commonly referred to as data assimilation [20].

Data assimilation for state updating aims at correcting for random error, but most hydrological data assimilation systems suffer from autocorrelated error, or bias, in either the observations or model simulations. A first methodology to remove the systematic mismatch between the model results and the observations was developed by Reichle and Koster [25]. In this algorithm, the external observations are rescaled, in order to match their Cumulative Density Function (CDF) to the CDF of the model forecasts. In other words, this method applies a static rescaling of the external data, and aims at assimilating anomaly information into the model. A different way of dealing with the bias issue is to modify the data assimilation algorithm, so the biases (forecast and/or observation) are estimated dynamically, in addition to the model state variables. This approach, focusing on forecast bias, was first introduced in meteorology [6, 7], and was then applied in land surface models [2, 12, 4, 3, 26]. Other studies have focused on the estimation of observation biases in addition to the model state variables. In meteorology, examples are Derber and Wu [9], Auligné et al. [1], Dee and Uppsala [8], and Fertig et al. [17]. Examples in hydrology that focus on the removal of remotely sensed soil moisture and temperature biases in a data assimilation framework can be found in Montzka et al. [23] and Draper et al. [11].

In the data assimilation algorithm that is arguably most frequently used in hydrology, the Kalman filter, two approaches have been used to estimate forecast biases in addition to the state variables. More specifically, the state vector can be augmented to include the bias estimates, as is ap-

plied in Drécourt et al. [12] and Kollat et al. [19]. A different approach is to apply the Kalman filter for the state estimates, and to run a second Kalman filter in parallel for the bias estimates [12, 4].

All previously mentioned studies have focused on estimating either the forecast bias, the observation bias, or a lumped observation and forecast bias in addition to the state variables. A first attempt to estimate both these biases individually and simultaneously in a synthetic framework was proposed by Pauwels et al. [24] for the assimilation of discharge into a conceptual hydrologic model.

A common feature of all studies focusing on separate bias and state variable estimation is a simplified propagation of the forecast bias and its error covariance. More specifically, the bias is often assumed to be persistent and the error covariance is calculated as a fraction of the a priori error covariance. This is a simplification of the explicit temporal propagation of the error covariances. The impact of this inconsistency, and the general lack of estimates for bias error covariances in general, on the performance of the filter is still not understood.

In this paper, we optimize the magnitude of the forecast bias error variances and we evaluate the impact of the choice of error propagations in a two-stage state and bias estimation method. The parameters of the bias estimation algorithm are estimated through a combination of global optimization and data assimilation. The method is applied in a synthetic assimilation case where groundwater storage is assimilated to improve discharge. For future applications, remotely sensed or in situ groundwater storage observations could be assimilated [13, 31, 10].

2. Site and Data Description

The study is performed in the Zwalm catchment in Belgium. Troch et al. [29] provide a complete description of this test site; only a very short overview is given here. The total drainage area of the catchment is 114.3 km² and the total length of the perennial channels is 177 km. The maximum elevation difference is 150 m. The average annual temperature is 10 °C, with January the coldest month (mean temperature 3 °C) and July the warmest month (mean temperature 18 °C). The average annual rainfall is 775 mm and is distributed evenly throughout the year. The annual actual evapotranspiration is approximately 450 mm.

Meteorological forcing data with a one-day time step from 1994 through 2002 were used in this study. The precipitation and all the variables needed to calculate the potential evapotranspiration using the Penman-Monteith equation were measured at the meteorological station in Kruishoutem. Discharge was measured continuously at the catchment outlet in Nederzwalm with an hourly time step, and daily averaged observations are derived for validation purposes.

3. Model Description

The Hydrologiska Byråns Vattenbalansavdelning (HBV) model, of which Figure 1 shows a schematic, was originally developed by Linström et al. [21]. In this paper, the version of Matgen et al. [22] is applied. The model uses observed precipitation ($R_{tot}(t)$) and potential evapotranspiration ($ETP(t)$) as input, both in ms^{-1} . t is the time in seconds. These are the only meteorological data sets needed to apply the model. Discharge is needed for calibration and validation, and an estimate of the catchment size is also needed to make the units of the modeled discharge consistent with the observations. The soil water in a catchment is divided into a soil reservoir ($S(t)$, m), a slow reservoir ($S_1(t)$, m), and a fast reservoir ($S_2(t)$, m). These reservoirs constitute the three prognostic variables of the system.

A number of fluxes are calculated, which depend on the state variables of the system. The actual evapotranspiration $ETR(t)$ (m^3s^{-1}) is first determined:

$$ETR(t) = \frac{1}{\lambda} \frac{S(t)}{S_{max}} ETP(t) \quad (1)$$

λ is a dimensionless parameter, and S_{max} is the storage capacity of the soil reservoir (m). The infiltration $R_{in}(t)$ (ms^{-1}) is calculated as follows:

$$R_{in}(t) = \left(1 - \frac{S(t)}{S_{max}}\right)^b R_{tot}(t) \quad (2)$$

b is a dimensionless parameter. After this, the effective precipitation $R_{eff}(t)$ (ms^{-1}) is determined:

$$R_{eff}(t) = R_{tot}(t) - R_{in}(t) \quad (3)$$

The percolation $D(t)$ (ms^{-1}) is then calculated:

$$D(t) = Pe \left(1 - e^{-\beta \frac{S(t)}{S_{max}}} \right) \quad (4)$$

Pe is a percolation parameter (ms^{-1}), and β a dimensionless parameter. After this, the storage in the soil reservoir at the end of the time step can be calculated as follows:

$$S(t + \Delta t) = S(t) + (R_{in}(t) - ETR(t) - D(t)) \Delta t \quad (5)$$

Δt is the time step in seconds. All three storage values are always positive.

The input in the fast reservoir $R_2(t)$ (ms^{-1}) then becomes:

$$R_2(t) = \alpha \frac{S(t)}{S_{max}} R_{eff}(t) \quad (6)$$

α is a dimensionless parameter. The outflow from this reservoir $q_2(t)$ (m^3s^{-1}) is then determined:

$$q_2(t) = \kappa_2 \left(\frac{S_2(t)}{S_{2,max}} \right)^\psi \quad (7)$$

$S_{2,max}$ is the storage capacity of the fast reservoir (m), and κ_2 (m^3s^{-1}) and ψ (-) are model parameters. After this, the storage in the fast reservoir at the end of the time step can be calculated as:

$$S_2(t + \Delta t) = S_2(t) + (R_2(t) - q_2(t)) \Delta t \quad (8)$$

The input in the slow reservoir $R_1(t)$ (ms^{-1}) becomes::

$$R_1(t) = R_{eff}(t) - R_2(t) \quad (9)$$

The outflow from this reservoir $q_1(t)$ (m^3s^{-1}) can be calculated as:

$$q_1(t) = \kappa_1 S_1(t) \quad (10)$$

κ_1 is a model parameter (m^2s^{-1}). Finally, the storage in the slow reservoir at the end of the time step is calculated:

$$S_1(t + \Delta t) = S_1(t) + (R_1(t) - q_1(t) + D(t)) \Delta t \quad (11)$$

The total discharge $q(t)$ is simply the sum of $q_1(t)$ and $q_2(t)$. A triangular unit hydrograph is used for runoff routing. Since in this paper daily time steps are used, and the concentration time of the catchment is only 14 hours [16], no routing needs to be performed for this study.

As a summary, the model contains ten time-invariant parameters (λ , S_{max} , b , α , P_e , β , ψ , $S_{2,max}$, κ_2 , and κ_1), and three state variables per time step (the storages $S(t)$, $S_1(t)$, and $S_2(t)$).

For reasons of computational efficiency, we have used this relatively simple model instead of a fully physical process-based land surface model. However, the conclusions drawn from this study can be expected to be valid for any model.

4. The Data Assimilation Algorithm

4.1. System Description

We will use the same system description as in De Lannoy et al. [4]. The system state is propagated from time step $k-1$ to time step k :

$$\tilde{\mathbf{x}}_k = f_{k-1,k}(\tilde{\mathbf{x}}_{k-1}, \mathbf{u}_k, \mathbf{w}_{k-1}) \quad (12)$$

$\tilde{\mathbf{x}}_k$ is the biased state vector, and \mathbf{u}_k is the vector with model inputs (for example the meteorologic forcing data). \mathbf{w}_{k-1} is the model error, which is a random error term with covariance \mathbf{Q}_{k-1} . $f_{k-1,k}$ is a nonlinear operator, propagating states and forcings at time step $k-1$ to states at time k . For the remainder of the paper, variables indicated with a $[\tilde{\cdot}]$ refer to biased variables, while variables without the $[\tilde{\cdot}]$ refer to unbiased variables. The difference between this equation and the equation used in the derivation of the ensemble Kalman filter [15] is that here biased state variables are used, while the ensemble Kalman filter is derived using unbiased variables.

The true unbiased state vector \mathbf{x}_k is defined as:

$$\mathbf{x}_k = \tilde{\mathbf{x}}_k - \mathbf{b}_k^m \quad (13)$$

\mathbf{b}_k^m is the true forecast bias. The system is observed as follows:

$$\mathbf{y}_k = h_k(\mathbf{x}_k, \mathbf{v}_k) \quad (14)$$

\mathbf{y}_k is the vector with the observations. \mathbf{v}_k is the zero mean observation error with covariance \mathbf{R}_k , and $h_k(\cdot)$ is the nonlinear observation operator.

The bias vector is propagated in two different ways. Following De Lannoy et al. [4], the true biases are persistent when moving from time step $k - 1$ to time step k :

$$\mathbf{b}_k^m = \mathbf{b}_{k-1}^m + \mathbf{w}_{k-1}^m \quad (15)$$

\mathbf{w}_{k-1}^m is the bias model error, which is a random error term with covariance \mathbf{Q}_{k-1}^m . In order to be able to model the bias evolution in a more general way, the following linear bias model will be tested:

$$\mathbf{b}_k^m = \mathbf{C}\mathbf{b}_{k-1}^m + \mathbf{d} + \mathbf{w}_{k-1}^m \quad (16)$$

\mathbf{C} is assumed to be a diagonal matrix, of which the entries need to be calibrated. \mathbf{d} is a vector, of which the entries need to be calibrated as well. The exact dimensions of the matrices and vectors will be discussed below (Section 5). It should be noted that a value of zero for \mathbf{d} and a unity value for \mathbf{C} turn this bias model into an autoregressive process.

In the following derivation of the filtering algorithm, it is important to stress that the errors in the forecast biases are assumed independent of the errors in the unbiased state of the system.

4.2. Propagation of the States and Biases

Figure 2 shows an overview of the equations used in the data assimilation framework. In a simulation framework, an estimate of the system state is used in the model. For the remainder of this paper, the notation $[\hat{\cdot}]$ refers to estimated variables. $[\hat{\cdot}]^-$ indicates an a priori estimate (forecast, before the update), and $[\hat{\cdot}]^+$ an a posteriori estimate (analysis, after the update). The state model is applied in an ensemble framework, with an ensemble size of N . First, the a posteriori state estimate is propagated:

$$\hat{\mathbf{x}}_k^{i-} = f_{k-1,k} \left(\hat{\mathbf{x}}_{k-1}^{i+}, \mathbf{u}_k^i, \mathbf{w}_{k-1}^i \right) \quad (17)$$

i indicates the ensemble member. The filter can be defined both with and without feedback [6, 12], but in this paper the bias is not fed back into the model. The system is observed as follows:

$$\hat{\mathbf{y}}_k^{i-} = h_k \left(\hat{\mathbf{x}}_k^{i-}, \mathbf{0} \right), \quad (18)$$

where $\hat{\mathbf{x}}_k^{i-}$ is an unbiased state estimate. The persistent biases are propagated in a deterministic way as follows:

$$\hat{\mathbf{b}}_k^{m-} = \hat{\mathbf{b}}_{k-1}^{m+} \quad (19)$$

In the case of a more general linear bias model, the biases are propagated by:

$$\hat{\mathbf{b}}_k^{m-} = \mathbf{C}\hat{\mathbf{b}}_{k-1}^{m+} + \mathbf{d} \quad (20)$$

The propagation of the state and bias is summarized in the first step of Figure 2. Figure 2 also shows that the bias estimates are not fed back into the model. To present and validate the results, the forecasted bias estimates are subtracted from the forecasted biased state estimates, i.e.

$$\hat{\mathbf{x}}_k^{i-} = \hat{\mathbf{x}}_k^{i-} - \hat{\mathbf{b}}_k^{m-} \quad (21)$$

4.3. Propagation of the Error Covariances

The second step in Figure 2 shows the calculation of the state and bias error covariances. The left hand boxes refer to the standard practice of assuming that the biased state error covariance and the forecast bias error covariance are a fraction of the unbiased state error covariance. This method was proposed by Dee [5] to simplify the calculation of the Kalman gain for the bias estimation. The parameter required for this, γ , is determined through calibration of the assimilation algorithm, as explained in Section 6.

The right hand side boxes of Figure 2 refer to an explicit propagation of the forecast bias error covariance. Appendix A presents the derivation of this bias error covariance propagation formulation. In this case the entries of \mathbf{C} and \mathbf{Q}^m need to be determined, as well as the entries in the vector \mathbf{d} . The advantage of this approach is clearly that the forecast bias error covariance can evolve independently of the state error covariance, and can be optimized independently.

4.4. State and Bias Updating

The remainder of the algorithm then consists of calculating the gain for the state and bias variables, updating the state and bias error covariance using these gains, and updating the state and bias estimates. Finally, the state estimates (ground water storage) are corrected for bias to present and validate the results. Yet, the bias correction is not fed back into the model, meaning that the model does not propagate the unbiased states, but rather the biased states. The discharge is modeled using the biased state and used as a direct model output in the validation. The discharge is produced with a biased system, i.e. a biased state yields an unbiased discharge by design. It should be noted that the bias update uses a Discrete Kalman filter (developed for linear systems), whereas the state update is performed with an ensemble Kalman filter.

5. Application of the Data Assimilation Algorithms

As explained in Section 3, there are three prognostic state variables and one observation. We are using observations of the ground water reservoir (S_1) to update the model. This is the part of the soil profile below the point where the pore pressure equals the atmospheric pressure down to the impermeable substrate. This can be measured, for example, by piezometers. We use a direct observation of the state variable. The main reason for assimilating the ground water storage is that the ground water storage has a much stronger impact on the discharge than the surface storage. Discharge observations will be used as an independent variable for model validation. An example of a study focusing on the assimilation of remotely sensed ground water storage values, and examining the impact on the modeled discharge, can be found in [31].

The observation system can be written as:

$$\mathbf{x}_k = \begin{bmatrix} S \\ S_1 \\ S_2 \end{bmatrix}_k, \quad \mathbf{y}_k = [S_1]_k + \mathbf{v}_k, \quad \mathbf{H} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \quad (22)$$

using a linear, time-invariant observation operator \mathbf{H} . Prior to performing the study, the observability of the system was checked, through calculating the rank of the observability matrix. This is written as:

$$\mathbf{O}_k = \begin{bmatrix} \mathbf{H} \\ \mathbf{H}\mathbf{F}_k \\ \mathbf{H}\mathbf{F}_k^2 \\ \dots \\ \mathbf{H}\mathbf{F}_k^{n-1} \end{bmatrix} \quad (23)$$

n is the number of state variables, and \mathbf{F}_k is the Jacobian matrix, resulting from the linearization of $f_{k-1,k}$ in Equation 12. This rank was found to be lower than three, which means that the values of the other state variables cannot be inferred from S_1 alone. Under the present setup, the only variable that can be estimated directly from observations of S_1 is S_1 itself. This can be explained by the model structure: S_1 receives water directly through re-infiltration of surface runoff, and indirectly through drainage from S . A certain value of S_1 can thus be obtained using different values of S and S_2 . This implies that one can never know what the values of S and S_2 are given

an observation of S_1 . For this reason, the remainder of the study is focusing on the estimation of the ground water storage S_1 (scalar), its forecast bias b^m (scalar) and the corresponding state and bias error variances (scalars).

6. Bias Model and Error Variance Estimation

6.1. General Framework

Throughout the study, the hydrologic model is applied in an ensemble mode, with 64 members. Pauwels et al. [24] have performed a sensitivity analysis with respect to the ensemble size for this model, and concluded that there was no benefit in using a higher number of ensemble members. Unless stated differently, the ensemble was generated by adding a Gaussian random number to all ten parameter values. The mean of this random number is zero, and the standard deviation is equal to 0.1 times the parameter value. It was ensured that non-negative values did not occur by setting lower limits on the obtained values. This is a realistic approach, and since the study is synthetic, similar conclusions would have been obtained using a different value.

A first step in the study consisted of calibrating the model using the observed discharge records at Nederzwalm, in order to obtain realistic parameter values. The Root Mean Square Error (RMSE) between the daily simulated and observed discharge was minimized, using Particle Swarm Optimization (see Section 6.4). Table 1 provides an overview of the obtained parameter values.

A synthetic experiment was then performed. The model simulations for discharge q and ground water storage (S_1) obtained using the parameters in Table 1 were assumed to be the synthetic truth. The synthetic observations (S_1^o and q^o) were generated by adding a random number with mean zero and standard deviation 0.1 mm to the synthetic true S_1 . Next, a biased truth was generated by adding a predefined bias of 2.5 mm to the true S_1 values (referred to as the increased \tilde{S}_1 values). These biased values are then used to calibrate the rainfall-runoff model with a forecast bias, i.e. the model is recalibrated, this time by minimizing the difference between the simulated and increased \tilde{S}_1 values, without any constraint on the discharge. The resulting biased parameter values are listed in Table 1. We acknowledge that bias can originate from biased forcings or initial conditions as well. However, the impact of the initial conditions would be limited in time as any model eventually converges to its climatology. The biased rainfall-runoff model was

then applied with explicit random forecast error introduced by perturbing the precipitation with a random number with mean zero and standard deviation 10% of the observed precipitation, ensuring that negative values were never obtained. The synthetically observed S_1^o values (free of bias, with random error) were then assimilated into the model (with bias and random error). These results were then evaluated by comparing the data assimilation results against the synthetic truth.

Three sets of data assimilation experiments were investigated in this Observation System Simulation Experiment (OSSE) framework. A first set involves experiments using a two-stage state and bias estimation with calibration of various filter parameters. In addition to the bias-aware Kalman filters, two other data assimilation strategies were also applied. On the one hand, the synthetic observations (S_1^o) were assimilated into the model using a bias-unaware Kalman filter. No filter parameters are estimated for this purpose and the forecast error is generated using the same spread multiplication factor (the random number with mean zero and standard deviation 10%, as defined above) as found after calibration of the bias-aware filter. On the other hand, the synthetic observations were rescaled so their CDF matches the CDF of the model simulations. The analysis of the results from these two algorithms allows us to demonstrate the importance of including bias estimates in the data assimilation framework.

6.2. The Calibration Set-Up

The calibration of the bias-aware filter experiments involves the estimation of (i) the bias model parameters (the diagonal values of the matrix \mathbf{C} , and the entries in the matrix \mathbf{d} , here reduced to scalars), (ii) the bias model error covariance \mathbf{Q}^m (here reduced to a scalar variance), or (iii) the filter parameter γ . All these parameters were assumed constant for the entire simulation period and were determined by minimizing an objective function. This objective function only aims at optimal estimation of the error variances in the Kalman filter, which can be measured by the Gaussianity of the innovations. More specifically, the normalized innovations for both the bias and state update equations must be Gaussian, i.e.

$$\frac{\mathbf{y}_k - \mathbf{H}_k \left(\hat{\mathbf{x}}_k^{i-} - \hat{\mathbf{b}}_k^{m-} \right)_{i=1,\dots,N}}{\sqrt{\mathbf{H}_k \left(\tilde{\mathbf{P}}_k^- + \mathbf{P}_k^{m-} \right) \mathbf{H}_k^T + \mathbf{R}_k}} \approx \mathcal{N}(0, 1), \quad \frac{\mathbf{y}_k - \left(\mathbf{H}_k \hat{\mathbf{x}}_k^{i-} + \mathbf{v}_k^i \right)_{i=1,\dots,N}}{\sqrt{\mathbf{H}_k \tilde{\mathbf{P}}_k^- \mathbf{H}_k^T + \mathbf{R}_k}} \approx \mathcal{N}(., 1), \quad (24)$$

respectively. $\tilde{\mathbf{P}}_k^-$ and \mathbf{P}_k^{m-} are the biased background error covariance and the bias error covariance, respectively. The overline indicates the ensemble average, and \mathbf{v}_k^i is an ensemble realization of the observation error (note again that the state update uses an ensemble Kalman filter). The mean of the first expression must be zero, and the standard deviation of both expressions must be equal to one. For this reason, two objective functions were defined, one for the bias, and one for the state variable, respectively. These objective functions were calculated by making a Gaussian probability plot with the ranked innovations (X) as abscissa and the corresponding Z -value as ordinate. The average of X (\overline{X}), which should be zero for the bias update innovations, the correlation between X and Y (R , which should be one), and the variance of X (σ_X^2 , which should also be one), were used for this purpose. For the bias, the objective function is written as:

$$\text{OF}_b = (1 - R)^2 + \overline{X}^2 + (1 - \sigma_X^2)^2 \quad (25)$$

For the state variable, the objective function is:

$$\text{OF}_s = (1 - R)^2 + (1 - \sigma_X^2)^2 \quad (26)$$

Because the magnitude of these two objective functions is similar, the simple sum of these two objective functions was minimized:

$$\text{OF} = \text{OF}_s + \text{OF}_b \quad (27)$$

This minimization was performed in two different calibration set-ups. A first calibration set-up was merely exploratory and focused purely on the estimation of the forecast bias model and the bias model error variance or the filter parameter γ , without conducting an independent validation. For this purpose the objective functions were calculated for the entire simulation period. In this experiment, synthetic observations were assimilated into the model with an assimilation interval of 7 days.

A second calibration set-up performed a separate calibration and validation of the bias model using the optimal calibration case identified during the exploratory calibration set-up. The calibration period consisted of the first three years of the simulation period. Using the obtained bias model, error covariances or filter parameter, the assimilation was applied for the remaining six years of the simulation period, during which the model results were validated. In this experiment the synthetic observations S_1^o were assumed

to be available with an assimilation interval of 3 days. For both calibration set-ups the parameter values, disturbances, and observations were the same.

In all set-ups, the initial bias estimates \mathbf{b}_0^m were zero, and the initial bias error variance (\mathbf{P}_0^{m-}) was set to 10 m^2 . Consequently, at the first time steps, the filter will practically replace the bias estimate with the initial innovation.

6.3. The Calibration Experiments

The bias model (\mathbf{C} and \mathbf{d}), the bias model error variance (\mathbf{Q}^m) and the filter parameter γ were estimated through the minimization of an objective function (Equation 27). Two extra calibration experiments were performed.

A first issue that was examined was whether or not it is beneficial to calibrate the ensemble state forecast spread in addition to the bias models. As stated in Section 6.1, the model parameters were perturbed by adding a random number with mean zero and a standard deviation of 0.1 times the parameter value. In this experiment, this multiplication factor was also calibrated, and the impact on the obtained results was assessed.

Secondly, we also checked whether it is necessary to have nonzero intercepts (\mathbf{d}) in the bias models. In this experiment, \mathbf{d} was set to zero, and only \mathbf{C} was calibrated.

6.4. The Global Optimization Algorithm

The parameter estimation algorithm used in this paper, particle swarm optimization (PSO), is based on the complex, collective behavior of individuals in decentralized, self-organizing systems. These systems are created through a population of individuals that interact locally with each other and with the community. These interactions lead to global behavior, which can result in the achievement of certain objectives. Examples of such systems in nature are abundant: ant colonies, swarms of birds, schools of fish, etc. [18]. For a complete description of the algorithm we refer to Scheerlinck et al. [27].

The results of PSO depend on the choice of several parameters: the population size N_p , the cognitive parameter c_1 , the social parameter c_2 , and the inertia weight w . Engelbrecht [14] found that a good value for the population size N_p is 30. In order to determine good values for the parameters c_1 , c_2 , and w , an exhaustive search was performed when calibrating the model using discharge observations. 36 Iterations were applied. The parameter c_1 was varied between 0.8 and 1.7, c_2 between 1 and 2.1, and w between 0.2 and 0.5, with steps of 0.1. The parameter values were chosen in the convergence

domain of particle swarm optimization, i.e., the region for which the population will converge [28]. 16 repetitions of the algorithm were performed, and the parameter combination leading to the lowest OF over the repetitions was retained. This has led to a value of 1.4 for c_1 , 2 for c_2 , and 0.4 for w . These values were used for the entire study.

7. Results

7.1. Calibration of the Bias Model

Table 2 shows the results of the calibration of the bias models for calibration set-up one. It should be stressed that the RMSE values in this table are calculated between the synthetic truth and the unbiased estimates of S_1 or q , while the objective function minimizes the difference between the ranked innovations and the corresponding Z -values of the Gaussian probability plot. For a general linear bias model, it is clear that calibrating the ensemble spread, and setting the intercept \mathbf{d} of the bias models to zero, leads to the best results. Leaving \mathbf{d} as a freely tunable parameter adversely impacts the RMSE performance in S_1 and q . This can be explained by the extra tunable parameter, which will ensure Gaussian innovations, but will not necessarily lead to a good estimation of the forecast bias. For persistent biases, calibrating the ensemble spread leads to the best simulations as well. Under these ideal calibration scenarios both forecast bias models lead to very similar RMSE values for S_1 and q , with S_1 slightly better estimated when using linear bias models. Based on these findings, the results obtained with a calibration of the ensemble spread, and for linear bias models with a zero intercept, will be retained for further analysis.

Table 2 shows that in these two cases the average biased state error variance is more than an order of magnitude larger than the average forecast bias error variance. This can be explained by the relatively low value of \mathbf{Q}^m and the relatively high value of γ , for the linear and persistent bias models, respectively. Step 2 in Figure 2 shows that in both cases this will lead to a low value of the forecast bias error variance. Table 2 also shows that in both these cases the spread multiplication factor is relatively similar.

A next step in the analysis is to examine the resulting forecast bias models. For linear forecast bias models, the following bias model has been obtained:

$$\mathbf{b}_k^m = 1.000039\mathbf{b}_{k-1}^m \quad (28)$$

This model is thus practically identical to a persistent bias model. Since calibrated model parameters are prone to uncertainty as well [30], and examining the results in Table 2, it is difficult to justify the use of a linear bias model other than the persistent bias model. For this reason, the remainder of the paper will focus on the results obtained using persistent bias models (i.e. $\mathbf{C} = \mathbf{I}$, $\mathbf{d} = 0$), while exploring the two different approaches to how the bias error variance is calculated.

7.2. Forecast Bias Error Variance Estimation

The recommended strategy for the calibration of the filter parameters (persistent bias models, and calibrating the ensemble spread) as determined in Section 7.1 is further investigated, in order to determine the best method to estimate the forecast bias error variance. For this purpose, the results from calibration set-up 2 are analyzed (three-year calibration period and six-year validation period), for both an explicit propagation of the forecast bias error variance, and a calculation as a fraction of the unbiased state error variance (referred to as the simplified calculation). This section describes the results during the validation period.

The top left panel of Figure 3 shows how the baseline model simulation performs compared to the true ground water storage S_1 . The bias in S_1 is 2.56 mm and the RMSE is 2.59 mm. Consequently, the S_1 state only has a random error standard deviation of approximately 0.51 mm (unbiased RMSE, assuming a static bias). This figure also shows that the model merges to its climatology very quickly, which justifies the bypassing of a model initialization period.

For the bias-unaware Kalman filter in the top right panel of Figure 3, bimodal model simulations are obtained. This can be explained by the typical saw-tooth behavior of the bias-unaware Kalman filter when applied to a biased system, in which the model simulations approximate the observations immediately after the update, but then quickly return to their original climatology, which degrades the time series correlation. These results indicate the importance of a good online forecast bias estimation.

The bottom panels show the results using a two-stage state and bias filter with explicit propagation of the forecast bias error variance. The biased results from the bias-aware Kalman filter only contain the state update without a posteriori correction for the bias estimate and are very similar to the results of the baseline run. In the bottom right panel, the unbiased results

show an excellent agreement with the synthetic truth, with the bias reduced to 0.1 mm, the RMSE is 0.26 mm, and the time series correlation R is 0.9.

Figure 4 shows the same analysis as Figure 3, but for the simplified calculation of the forecast bias error variance. Overall, the same conclusions can be drawn as for Figure 3. The results of the bias-aware filter are very similar to the results obtained using the explicit calculation of the forecast bias error variance. It should be noted that the results of the bias-unaware filters are slightly different in both cases because of the different multiplication factors for the parameter perturbation, which are listed in Table 3.

Table 3 lists the values for the bias filter parameter, spread multiplication factor, the average biased state error variance and forecast model error variance, and the innovation statistics for the two error variance estimation methods. Similar as in Section 7.1, the spread multiplication factors (for the state perturbation) are relatively similar, and a low value for \mathbf{Q}^m and a high value for γ have been obtained. This leads to relatively low values for the forecast bias error variance (\mathbf{P}^m) as compared to the biased state error variance ($\tilde{\mathbf{P}}$). The random error in the biased state is thus larger than in the bias and consequently most of the assimilated information is used to correct the state. This is in line with the expectations of this synthetic experiment, where the state is explicitly perturbed with random noise in precipitation, whereas the uncertainty on the imposed bias is limited to random variations due to the generation of the forecast bias through the use of different model parameters. In both cases, the standard deviations of the the un-normalized innovations for the state variables and the bias are relatively similar, while the standard deviations of the normalized innovations are relatively close to 1. This indicates that the functioning of the filter fulfills the theoretical requirements.

Figure 5 shows the temporal evolution of the bias. The true bias is calculated as the difference between the model simulations obtained with the biased and the unbiased model parameters. Peaks in the true bias in Figure 5 may thus be random error instead of bias, but it is impossible to objectively divide the difference between these two model simulations into a random component and bias. The bias does not show a constant value of 2.5 mm, as was imposed to the biased observations to re-calibrate the biased model. Instead, there is a mild annual cycle, dominated by the annual cycle in the evapotranspiration.

The bias estimates with both the explicit and simplified propagation of the forecast bias error variance lead to a good match with the true forecast

bias. Figure 5 further shows that the explicit calculation of the forecast bias error variance leads to a faster convergence of the estimated bias to the true bias, and a lower temporal variability in the bias estimate, whereas the simplified calculation requires a longer spinup and better reflects the temporal pattern, but with a time lag.

Figure 6 further explains these results. Consistent with Table 3, the biased state error variances (top panel of Figure 6) obtained with both methods are very similar and determined by the spread multiplication factor used to perturb the biased state. The magnitude of $\tilde{\mathbf{P}}^-$ is correctly calibrated to the uncertainty introduced in the biased model forecasts after re-calibration and perturbation of the precipitation. This will lead to very similar values of the state variable gains (third plot). However, the explicitly propagated forecast bias error variance is initialized at a large value of 10 m^2 , whereas it equals $\tilde{\mathbf{P}}^-$ in the simplified calculation, i.e. when the forecast bias error variance is modeled as a fraction of the unbiased state error variance. This will lead to a forecast bias gain that is closer to -1 at the onset of the simulations with explicitly propagated bias error variances, when \mathbf{P}^m is still large. The absolute value of the bias gain rapidly reduces after each update (step 4 in Figure 2) and the low value for \mathbf{Q}^m causes the forecast bias error variance to stabilize at a low value with the bias Kalman gain close to 0. With the simplified calculation, the bias error variance is slightly larger and the bias Kalman gain has a slightly larger absolute value. Because the biased state error variance values are more than an order of magnitude larger than the forecast bias error variance values (except at the onset of the simulations for the explicit propagation), the magnitude of the state variable Kalman gains will be larger than the forecast bias Kalman gains, as discussed above. The conclusion that can be drawn from Figure 6 is that in both cases the state variable estimates have a higher weight in the filter than the bias estimates. The initialization of the bias error variances causes a faster convergence of the estimated bias to the truth for the explicit propagation, but a lower temporal variability as the simulation progresses.

Figure 7 shows the impact of the assimilation of the ground water storage on the modeled discharge, when explicitly propagating the forecast bias error variance. The top left panel exposes a clear dynamic bias (mean bias = $1.61 \text{ m}^3\text{s}^{-1}$) in the baseline simulation of discharge, with an RMSE of $1.71 \text{ m}^3\text{s}^{-1}$ and $R=0.81$. The bias in the modeled storage leads to a relatively high bias in the discharge, which can be partly undone (mean bias = $0.78 \text{ m}^3\text{s}^{-1}$) when assimilating bias-free S_1 observations with a bias-unaware Kalman filter (top

right), but at the expense of a degraded $R = 0.74$. By design, the discharge diagnosed using the biased S_1 (bottom left) in the bias-aware Kalman filter, remains biased because only random state error is corrected for. However, when comparing the synthetic truth to the discharge obtained after correcting S_1 with the dynamic bias estimate (bottom right), one can conclude that the bias in the discharge has practically disappeared (mean bias = $0.08 \text{ m}^3\text{s}^{-1}$), the RMSE is reduced to $0.64 \text{ m}^3\text{s}^{-1}$ and $R = 0.82$. The remaining errors can be explained by the use of different model parameters to establish the unbiased and the biased model simulations. Even if the unbiased state variables are well estimated (which is the case), the use of these different parameters will lead to discharge values that are different from the truth.

Figure 8 shows the same analysis as Figure 7 for the simplified calculation. Similar conclusions as for Figure 7 can be drawn. The difference is that the calibration resulted in a slightly higher $\tilde{\mathbf{P}}^-$ and \mathbf{P}^m , so that the updates to both the bias and the state estimate are slightly stronger.

The conclusion from this section is that the unbiased state and discharge estimates are very similar after calibration of the explicit and simplified propagation of the forecast bias error variances and that the calibration based on innovation statistics successfully estimated the relative uncertainty in the bias and state estimates.

7.3. Comparison to the Observations Rescaling Approach

The results from the bias-aware Kalman filter were also compared to the commonly applied approach of rescaling the observations to the climatology of the model [25]. For this reason, the synthetic S_1 observations, for both experiments, were rescaled, so their Continuous Distribution Function (CDF) matches the CDF of the model simulations. The synthetic observations are assimilated into the model with an assimilation frequency of 3 days, with a three year calibration period, and a six-year validation period. Here, only the spread multiplication factor for the state uncertainty is calibrated. The results during the validation period are analyzed.

Table 4 shows the results of assimilating these rescaled observations into the model. One can immediately see that these results are only marginally better than the results from the baseline run, in terms of both the RMSE and the bias. This is what can be expected, because rescaling the observations to the biased model climatology will not remove the forecast bias, and the resulting ground water storage and discharge will remain biased. However, the temporal variability should be improved. If the bias were static (as assumed

in CDF-matching), the unbiased RMSE between the discharge simulations and observations obtained through the assimilation of the CDF-matched observations should be lower than this value for the baseline run. The unbiased RMSE is calculated as:

$$\text{RMSE}_u = \sqrt{\text{RMSE}^2 - b^2} \quad (29)$$

b is the bias between the truth and the simulations. For the baseline and assimilation runs, this unbiased RMSE is 0.68 and 0.67 m^3s^{-1} , respectively. This means that the temporal evolution does show a very slight improvement. CDF-matching can improve the variability of the assimilated variable (in this case ground water storage) through data assimilation, but it does not correct the (possibly dynamic) bias. The logical consequence is that CDF-matching may not be adequate in coupled models, or for the improvement in diagnostic variables for which the correct absolute value of the storage is important, such as for example discharge, brightness temperature, or crop yield.

These results further prove the fact that the two methods to deal with bias (CDF-matching and online estimation) result in a similar, improved variability in the diagnostic variable (discharge), but that CDF matching cannot improve the unbiased value of this variable.

8. Conclusions

The objective of this paper was to examine to which extent the choice for the propagation of the bias estimate and its error covariance influence the performance of the bias-aware Kalman filter. This was performed through a combination of global optimization of tuning parameters and ensemble Kalman filtering. A synthetic experiment was developed, and a conceptual rainfall-runoff model was used for this purpose. Ground water storage observations were assimilated into the model. A bias-aware Kalman filter was applied for a specific calibration period, during which an objective function was optimized to estimate error variances and bias model parameters. This objective function consisted of the deviation of the ranked normalized innovations from the corresponding Z -values in a Gaussian probability plot, calculated for both the state variable and the bias updates. A persistent or a more general linear forecast bias model were tested to model the bias propagation, and the persistent model was retained, because the more general linear model was nearly identical to a persistent model after calibration.

Two approaches for calculating the forecast bias error covariance have been further investigated, more specifically through the calculation as a constant fraction of the unbiased state error covariance, or through explicit propagation. After calibration, both methods have been found to lead to a very similar model performance. The calibration successfully attributes realistic error variance estimates to the bias and the state. The variance magnitudes are in line with the imposed uncertainty magnitudes in the synthetic experiment. Further, the results obtained using the bias-aware Kalman filter clearly outperform the results from a bias-unaware filter. These results have also been compared to the results obtained with assimilation of observations after rescaling to the model climatology. When CDF-matching is applied, only a marginal improvement in the resulting discharge as compared to the baseline run has been obtained. CDF-matching thus helps to improve the variability in the diagnostic variable (discharge), but cannot help to improve the unbiased estimate.

To summarize, the results lead to the conclusion that incorporation of the bias estimation in the soil moisture assimilation algorithm is the preferred method when compared to CDF-matching, especially if there is an interest in diagnostic variables that depend on good estimates of actual soil moisture. The filter parameters can be successfully estimated by calibrating the Gaussianity of the normalized innovations, regardless of the choice for the propagation of the forecast bias error variance. Overall, the results in this paper justify the use of the traditional method for online bias estimation with a persistent bias model and the estimation of the forecast bias error covariance as a fraction of the unbiased state error covariance.

9. Acknowledgements

Valentijn Pauwels is currently funded by the Australian Research Council through the Future Fellowship scheme. We thank the Department Operational Water Management of the Flemish Environmental Agency (Belgium) for providing the discharge and the meteorological data.

References

- [1] Auligné, T., McNally, A. P., Dee, D. P., 2007. Adaptive bias correction for satellite data in a numerical weather prediction system. *Quart. J. Royal Meteor. Soc.* 133 (624), 631–642.

- [2] Baek, S.-J., Hunt, B. R., Kalnay, E., Ott, E., 2006. Local ensemble kalman filtering in the presence of model bias. *Tellus A* 51A, 293–306.
- [3] Bosilovich, M. G., Radakovich, J. D., da Silva, A., Todling, R., Verter, F., 2007. Skin temperature analysis and bias correction in a coupled land-atmosphere data assimilation system. *J. Meteor. Soc. Japan* 85A, 205–228.
- [4] De Lannoy, G. J. M., Houser, P. R., Pauwels, V. R. N., Verhoest, N. E. C., 2007. Correcting for forecast bias in soil moisture assimilation with the Ensemble Kalman Filter. *Water Resour. Res.* 43 (9).
- [5] Dee, D. P., 2005. Bias and data assimilation. *Quart. J. Royal Meteor. Soc.* 131, 3323–3343.
- [6] Dee, D. P., Da Silva, A. M., 1998. Data assimilation in the presence of forecast bias. *Quart. J. R. Meteorol. Soc.* 124 (545), 269–295.
- [7] Dee, D. P., Todling, R., 2000. Data assimilation in the presence of forecast bias: The geos moisture analysis. *Month. Weather Rev.* 128 (9), 3268–3282.
- [8] Dee, D. P., Uppsala, S., 2009. Variational bias correction of satellite radiance data in the ERA-interim reanalysis. *Quart. J. Royal Meteor. Soc.* 135 (644), 1830–1841.
- [9] Derber, J. C., Wu, W. S., 1998. The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Monthly Weather Rev.* 126 (8), 2287–2299.
- [10] Dijk, A. I. J. M. V., Renzullo, L. J., Wada, Y., Tregoning, P., 2014. A global water cycle reanalysis (2003-2012) merging satellite gravimetry and altimetry observations with a hydrological multi-model ensemble. *Hydrol. Earth Syst. Sci.* 8 (18), 2955–2973.
- [11] Draper, C., Reichle, R., De Lannoy, G., Scarino, B., 2014. A dynamic approach to addressing observation-minus-forecast mean differences in a land surface skin temperature assimilation system. *J. Hydromet.* in press.

- [12] Drécourt, J. P., Madsen, H., Rosbjerg, D., 2006. Bias aware Kalman filters: Comparison and improvements. *Adv. Water Resour.* 29 (5), 707–718.
- [13] Eigbe, U., Beck, M. B., Wheeler, H. S., Hirano, F., 1998. Kalman filtering in groundwater flow modelling: problems and prospects. *Stoch. Hydrol. and Hydraul.* 12 (1), 15–32.
- [14] Engelbrecht, A., 2003. *Fundamentals of Computational Swarm Intelligence*. John Wiley and Sons Ltd., London, UK.
- [15] Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* 99 (C5), 10143–10162.
- [16] Ferket, B. V. A., Samain, B., Pauwels, V. R. N., 2008. Internal validation of conceptual rainfall-runoff models using baseflow records. *J. Hydrol.* 381 (1-2), 158–173.
- [17] Fertig, E., Baek, S., Hunt, B. R., Ott, E., Szunyogh, I., Arevequia, J. A., Kalnay, E., Li, H., Liu, J., 2009. Observation bias correction with an ensemble kalman filter. *Tellus A* 61A, 210–226.
- [18] Kennedy, J., Eberhart, R. C., 1995. Particle Swarm Optimization. In: *Proc. IEEE In. Conf. Neur. Netw.* Piscataway, NJ, pp. 1942–1948.
- [19] Kollat, J. B., Reed, P. M., Maxwell, R. M., 2011. Many-objective groundwater monitoring network design using bias-aware ensemble Kalman filtering, evolutionary optimization, and visual analytics. *Water Resour. Res.* (47).
- [20] Kostov, K. G., Jackson, T. J., 1993. Estimating profile soil moisture from surface layer measurements - A review. In: *Proc. of the International Society of Optical Engineering (SPIE)*. Vol. 1941. Orlando, USA, pp. 125–136.
- [21] Linström, G., Johansson, B., Persson, M., Gardelin, M., Bergström, S., 1997. Development and test of the distributed HBV-96 hydrological model. *J. Hydrol.* 201 (1-4), 272–288.

- [22] Matgen, P., Henry, J., Hoffmann, L., Pfister, L., 2006. Assimilation of remotely sensed soil saturation levels in conceptual rainfall-runoff models. In: IAHS Red Book Series, Prediction in Ungauged Basins: Promise and Progress. Vol. IAHS Publ. 303.
- [23] Montzka, C., Grant, J., Moradkhani, H., Hendricks Franssen, H.-J., Weihermueller, L., Drusch, M., Vereecken, H., 2012. Estimation of radiative transfer parameters from SMOS brightness temperatures using data assimilation: Implication on soil moisture retrieval. *Vadose Zone Journal* 12 (3).
- [24] Pauwels, V. R. N., De Lannoy, G. J. M., Hendricks Franssen, H.-J., Vereecken, H., 2013. Simultaneous estimation of model state variables and observation and forecast biases using a two-stage hybrid kalman filter. *Hydrol. Earth Syst. Sci.* 17 (9), 4525–4540.
- [25] Reichle, R. H., Koster, R. D., 2004. Bias reduction in short records of satellite soil moisture. *Geophys. Res. Lett.* (31).
- [26] Reichle, R. H., Kumar, S. V., Mahanama, S. P. P., Koster, R. D., Liu, Q., 2010. Assimilation of satellite-derived skin temperature observations into land surface models. *J. Hydromet.* 11, 1103–1122.
- [27] Scheerlinck, K., Pauwels, V. R. N., Vernieuwe, H., De Baets, B., 2009. Calibration of a water and energy balance model: Recursive parameter estimation versus particle swarm optimization. *Water Resour. Res.* 10 (45).
- [28] Trelea, I. C., 2003. The particle swarm optimization algorithm: convergence analysis and parameter selection. *Inf. Process. Lett.* 85 (6), 317–325.
- [29] Troch, P. A., De Troch, F. P., Brutsaert, W., 1993. Effective water table depth to describe initial conditions prior to storm rainfall in humid regions. *Water Resour. Res.* 29 (2), 427–434.
- [30] Vrugt, J. A., Ter Braak, C. J. F., 2011. DREAM((D)): an adaptive Markov Chain Monte Carlo simulation algorithm to solve discrete, non-continuous, and combinatorial posterior parameter estimation problems. *Hydrol. Earth Syst. Sci.* 12 (15), 3701–3713.

- [31] Zaitchik, B. F., Rodell, M., Reichle, R. H., 2008. Assimilation of GRACE terrestrial water storage data into a Land Surface Model: Results for the Mississippi river basin. *J. Hydromet.* 9 (3), 535–548.

Table 1: The calibrated model parameters.

Parameter	Calibrated Value	Biased Value	Units
λ	0.4072	4.8757	-
S_{max}	0.8403	0.9741	m
b	0.6923	0.3406	-
α	0.9973	1.107	-
P_e	$2.8598 \cdot 10^{-7}$	$4.6298 \cdot 10^{-8}$	ms^{-1}
β	0.1672	1.6203	-
ψ	3.4903	1.6211	-
$S_{2,max}$	0.0100	0.1725	m
κ_2	$1.9903 \cdot 10^{-8}$	$1.7764 \cdot 10^{-6}$	ms^{-1}
κ_1	$9.2567 \cdot 10^{-6}$	$6.0608 \cdot 10^{-6}$	s^{-1}

Table 2: RMSE values obtained from the different calibration experiments for the 9-year calibration period. \mathbf{Q}^m is in mm, and $\tilde{\mathbf{P}}_k^-$ and $\overline{\mathbf{P}}_k^{m-}$ are in mm². P and L stand for a persistent and a linear bias model, respectively.

Bias model	Calibrating ensemble spread	Setting \mathbf{d} to zero	RMSE S_1 (mm)	RMSE q (m ³ s ⁻¹)	Filter Parameter	Spread multiplication factor	$\tilde{\mathbf{P}}_k^-$	$\overline{\mathbf{P}}_k^{m-}$
L	Y	Y	0.338	1.332	$\mathbf{Q}^m=1\cdot 10^{-6}$	0.109	$2.104\cdot 10^{-1}$	$7.161\cdot 10^{-3}$
L	N	Y	0.345	2.818	$\mathbf{Q}^m=1\cdot 10^{-6}$	0.1	4.262	$1.251\cdot 10^{-1}$
L	Y	N	140.225	99.192	$\mathbf{Q}^m=1\cdot 10^{-6}$	0.119	$2.511\cdot 10^{-1}$	$8.058\cdot 10^2$
L	N	N	1.982	3.385	$\mathbf{Q}^m=9.783\cdot 10^{-3}$	0.1	3.897	1.532
P	Y	Y	0.417	1.335	$\gamma=0.954$	0.116	$3.097\cdot 10^{-1}$	$1.480\cdot 10^{-2}$
P	N	Y	0.370	2.812	$\gamma=0.941$	0.1	5.440	$3.394\cdot 10^{-1}$

Table 3: Innovation statistics, average error variances, and obtained filter parameters during the validation period. Units for the error variances and for the standard deviations of the un-normalized innovations are mm^2 and mm , respectively. The top line presents results for the simplified error covariance estimation, the bottom line the results for the explicit propagation. The observation error variance is in all these experiments equal to 0.01 mm^2 .

Bias filter parameter	State spread multiplication factor	$\bar{\mathbf{P}}_k^-$	$\bar{\mathbf{P}}_k^{m-}$	Innovation standard deviation			
				Un-normalized		Normalized	
				Bias	State	Bias	State
$\mathbf{Q}^m=2.950 \cdot 10^{-6}$	0.149	$3.073 \cdot 10^{-1}$	$2.676 \cdot 10^{-3}$	$4.469 \cdot 10^{-1}$	$4.266 \cdot 10^{-1}$	$7.886 \cdot 10^{-1}$	$8.996 \cdot 10^{-1}$
$\gamma=0.975$	0.161	$3.628 \cdot 10^{-1}$	$9.486 \cdot 10^{-3}$	$3.275 \cdot 10^{-1}$	$3.891 \cdot 10^{-1}$	$6.048 \cdot 10^{-1}$	$9.539 \cdot 10^{-1}$

Table 4: Results of the comparison of the model results, obtained by assimilating the CDF-matched observations, to the unbiased truth. The truth is in abscissa, the model results in ordinate. Discharge is in m^3s^{-1} , S_1 is in mm. All statistics are calculated during the validation period, from 1997 through 2002.

Variable	Average Truth	Average Simulations	Slope	Intercept	R	RMSE
q	1.119	2.686	1.230	1.311	0.814	1.767
S_1	0.990	3.407	0.932	2.484	0.883	2.540

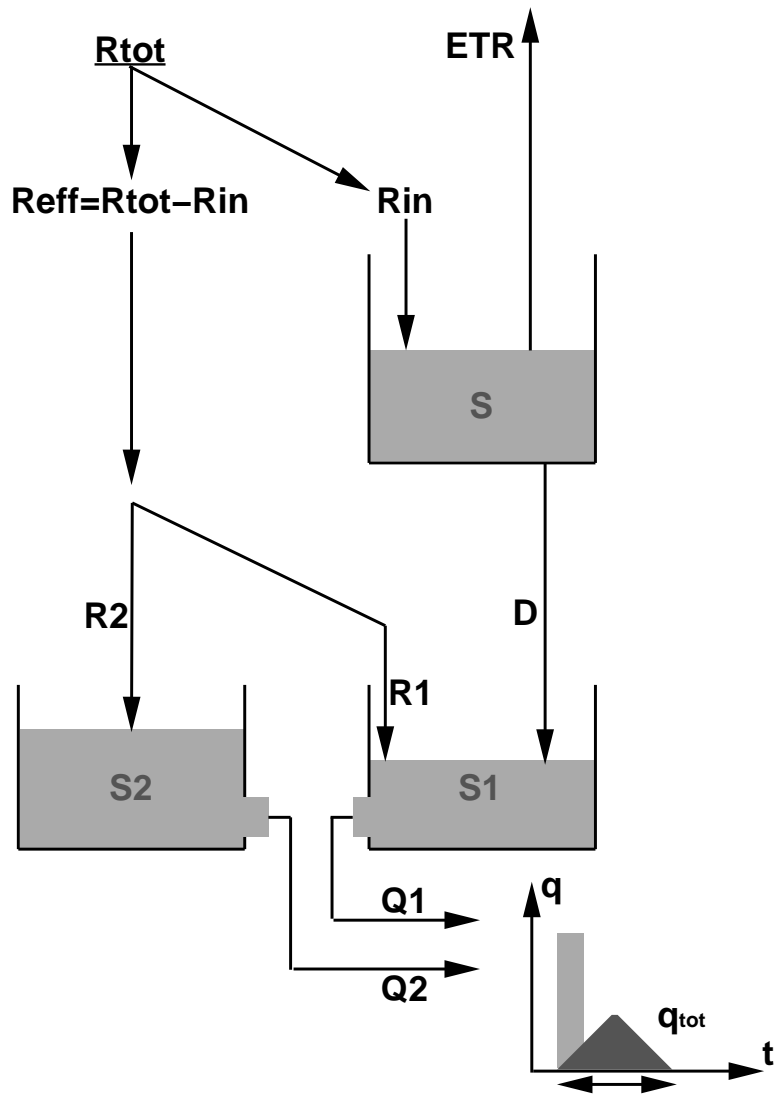


Figure 1: Schematic of the model.

1. Propagation of the state and bias estimates

<p>Persistent biases</p> $\begin{cases} \hat{\mathbf{x}}_k^{i-} &= f_{k-1,k}(\hat{\mathbf{x}}_{k-1}^{i+}, \mathbf{u}_k, \mathbf{w}_{k-1}^i) \\ \hat{\mathbf{b}}_k^{m-} &= \hat{\mathbf{b}}_{k-1}^{m+} \end{cases}$	<p>Linear bias models</p> $\begin{cases} \hat{\mathbf{x}}_k^{i-} &= f_{k-1,k}(\hat{\mathbf{x}}_{k-1}^{i+}, \mathbf{u}_k, \mathbf{w}_{k-1}^i) \\ \hat{\mathbf{b}}_k^{m-} &= \mathbf{C}\hat{\mathbf{b}}_{k-1}^{m+} + \mathbf{d} \end{cases}$
--	--

2. Estimation of the error covariances

$\begin{cases} \tilde{\mathbf{P}}_k^- &= \mathbb{E} \left[\left(\tilde{\mathbf{x}}_k - \hat{\mathbf{x}}_k^- \right) \left(\tilde{\mathbf{x}}_k - \hat{\mathbf{x}}_k^- \right)^T \right] \\ \tilde{\mathbf{P}}_k^- &= \gamma \mathbf{P}_k^- \\ \mathbf{P}_k^{m-} &= (1 - \gamma) \mathbf{P}_k^- \end{cases}$	$\begin{cases} \tilde{\mathbf{P}}_k^- &= \mathbb{E} \left[\left(\tilde{\mathbf{x}}_k - \hat{\mathbf{x}}_k^- \right) \left(\tilde{\mathbf{x}}_k - \hat{\mathbf{x}}_k^- \right)^T \right] \\ \mathbf{P}_k^- &= \tilde{\mathbf{P}}_k^- + \mathbf{P}_k^{m-} \\ \mathbf{P}_k^{m-} &= \mathbf{C}\mathbf{P}_{k-1}^{m+}\mathbf{C}^T + \mathbf{Q}^m \end{cases}$
--	--

3. Calculation of the bias Kalman Gain

$$\mathbf{K}_k^m = -\mathbf{P}_k^{m-} \mathbf{H}_k^T \left[\mathbf{H}_k \left(\tilde{\mathbf{P}}_k^- + \mathbf{P}_k^{m-} \right) \mathbf{H}_k^T + \mathbf{R}_k \right]^{-1}$$

4. Updating of the bias error covariance

$$\mathbf{P}_k^{m+} = [\mathbf{I} + \mathbf{K}_k^m \mathbf{H}_k] \mathbf{P}_k^{m-}$$

5. Calculation of the state Kalman Gain

$$\mathbf{K}_k = \tilde{\mathbf{P}}_k^- \mathbf{H}_k^T \left[\mathbf{H}_k \tilde{\mathbf{P}}_k^- \mathbf{H}_k^T + \mathbf{R}_k \right]^{-1}$$

6. Updating of the state error covariance

$$\tilde{\mathbf{P}}_k^+ = [\mathbf{I} - \mathbf{K}_k \mathbf{H}_k] \tilde{\mathbf{P}}_k^-$$

7. Updating of the bias estimate

$$\hat{\mathbf{b}}_k^{m+} = \hat{\mathbf{b}}_k^{m-} + \mathbf{K}_k^m \left(\mathbf{y}_k - \overline{\mathbf{H}_k \left(\hat{\mathbf{x}}_k^{i-} - \hat{\mathbf{b}}_k^{m-} \right)}_{i=1, \dots, N} \right)$$

8. Updating of the state estimates

<p>With bias correction</p> $\hat{\mathbf{x}}_k^{i+} = \hat{\mathbf{x}}_k^{i-} - \hat{\mathbf{b}}_k^{m+} + \mathbf{K}_k \left(\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^{i-} + \mathbf{H}_k \hat{\mathbf{b}}_k^{m+} + \mathbf{v}_k^i \right)$	<p>Without bias correction</p> $\hat{\mathbf{x}}_k^{i+} = \hat{\mathbf{x}}_k^{i-} + \mathbf{K}_k \left(\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^{i-} + \mathbf{v}_k^i \right)$
---	---

Figure 2: The data assimilation algorithms. All terms are explained in Section 4.1.

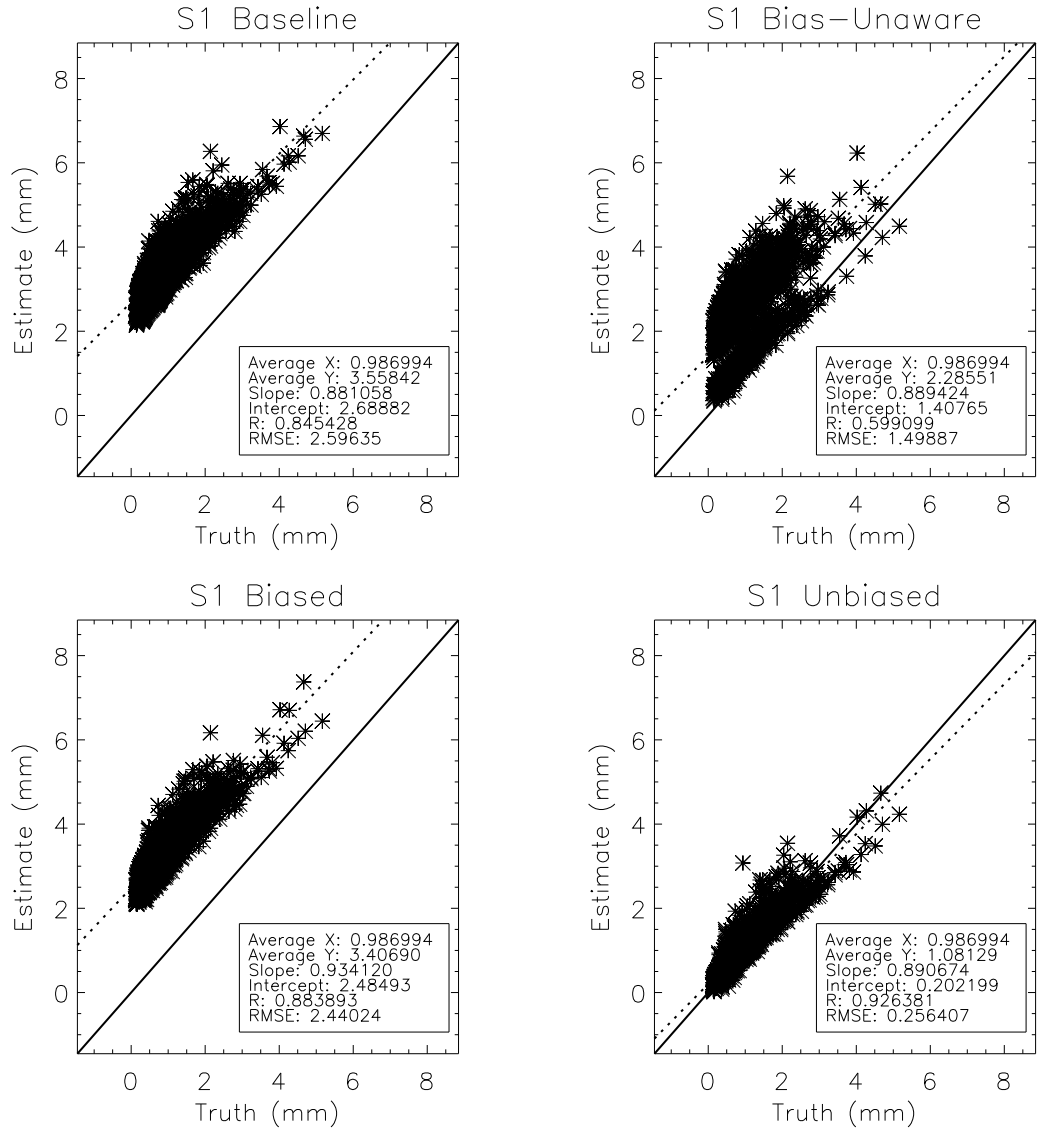


Figure 3: The modeled groundwater storage for the persistent forecast bias model with explicit propagation of the forecast bias error covariance. Data are shown for the 6-year validation period following a 3-year calibration period. The top left hand side panel shows the results of the baseline run and the top right hand side the results of the bias-unaware filter. The bottom panels show the results of the bias-aware filter, with the left hand side panel the biased and the right hand side panel the unbiased estimates.

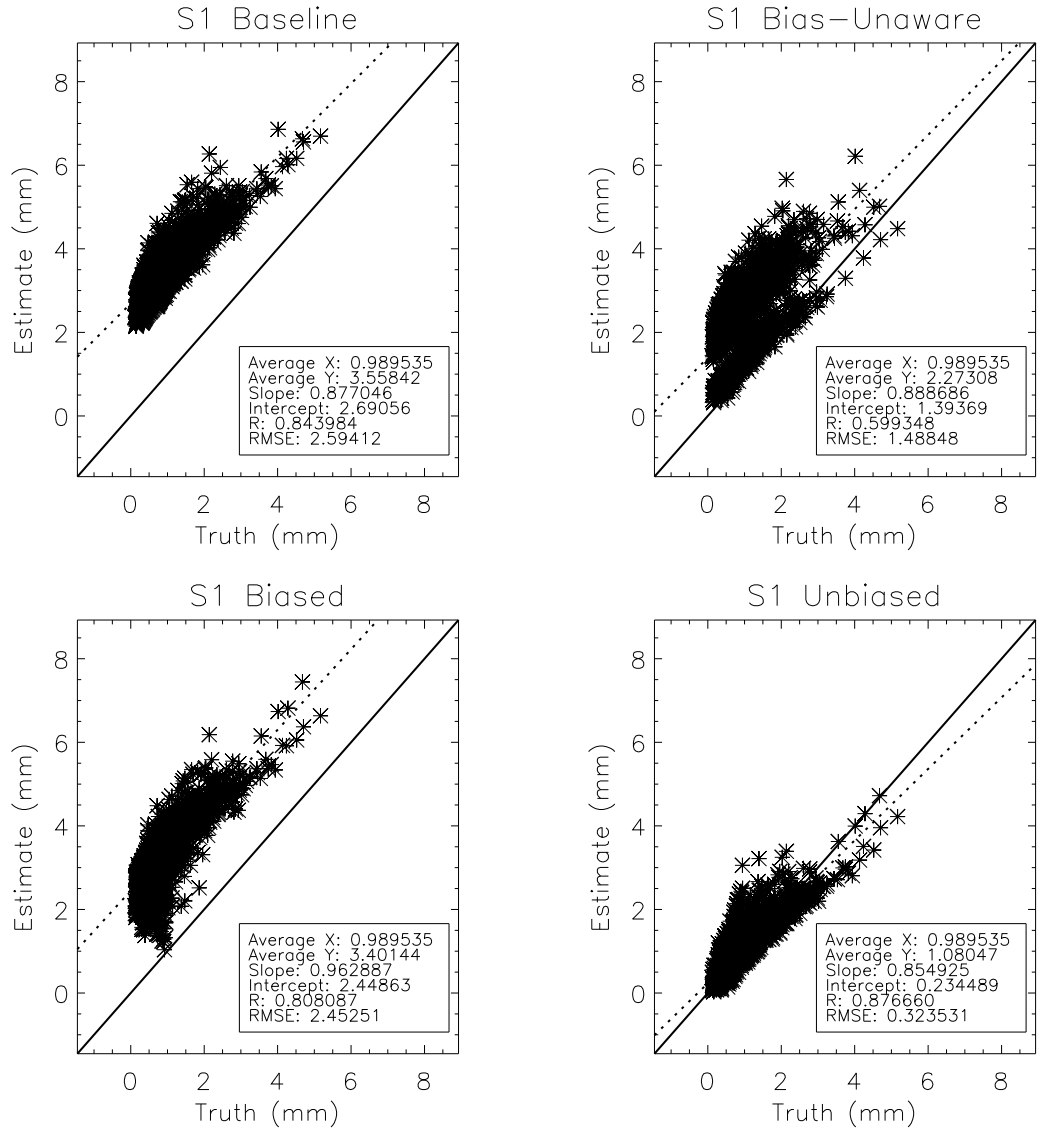


Figure 4: The modeled groundwater storage for the persistent forecast bias model with simplified calculation of the forecast bias error covariance. Data are shown for the 6-year validation period following a 3-year calibration period. The top left hand side panel shows the results of the baseline run and the top right hand side the results of the bias-unaware filter. The bottom panels show the results of the bias-aware filter, with the left hand side panel the biased and the right hand side panel the unbiased estimates.

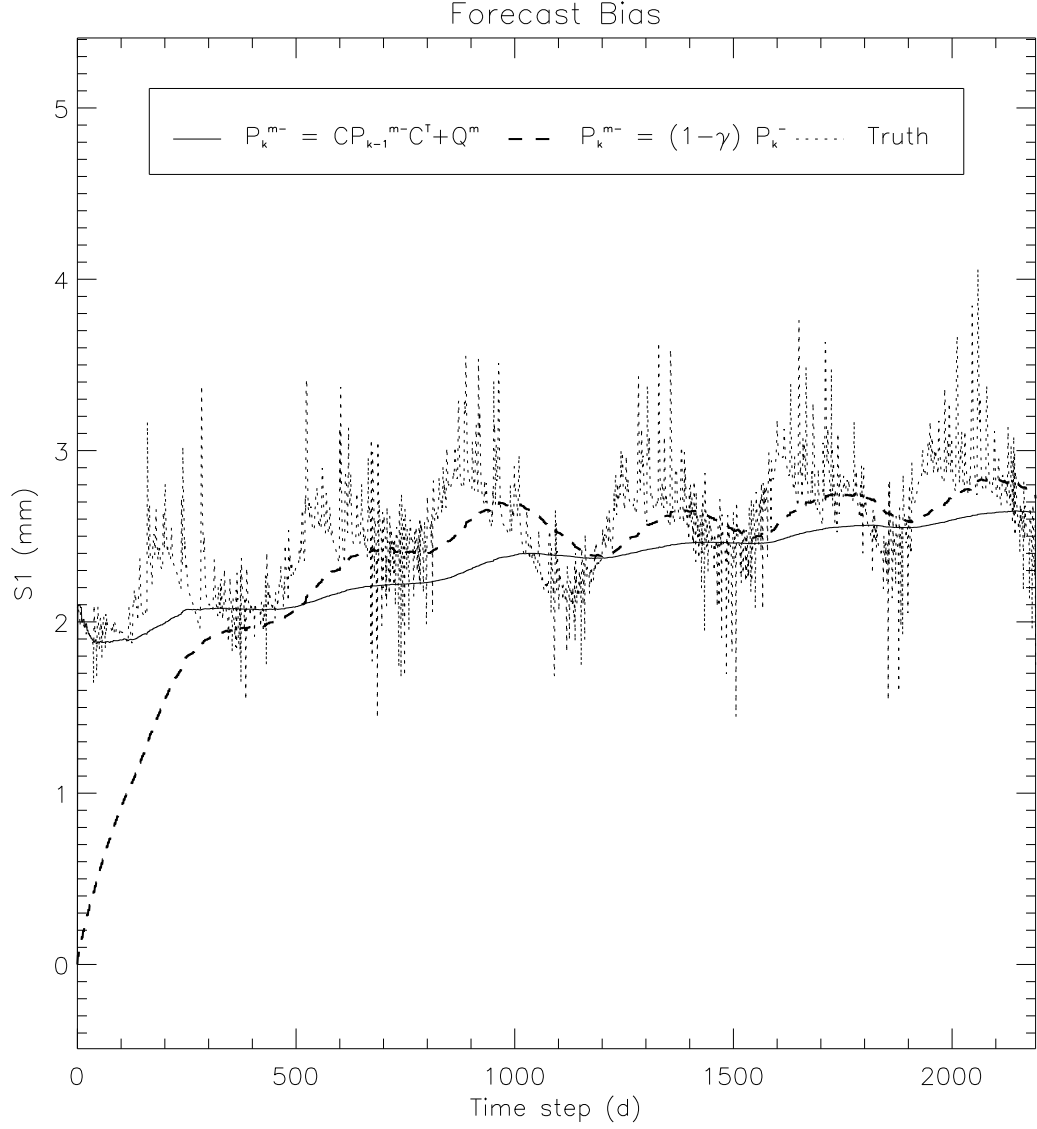


Figure 5: The modeled bias in the validation period, using a persistent forecast bias model and two approaches to propagate \mathbf{P}^m . The true bias is approximated as the difference in the model results obtained with the recalibrated and original model parameters. The dotted line represents the synthetic true S_1 , the solid line the estimated S_1 using explicit bias propagation, and the thick dashed line S_1 using the simplified calculation.

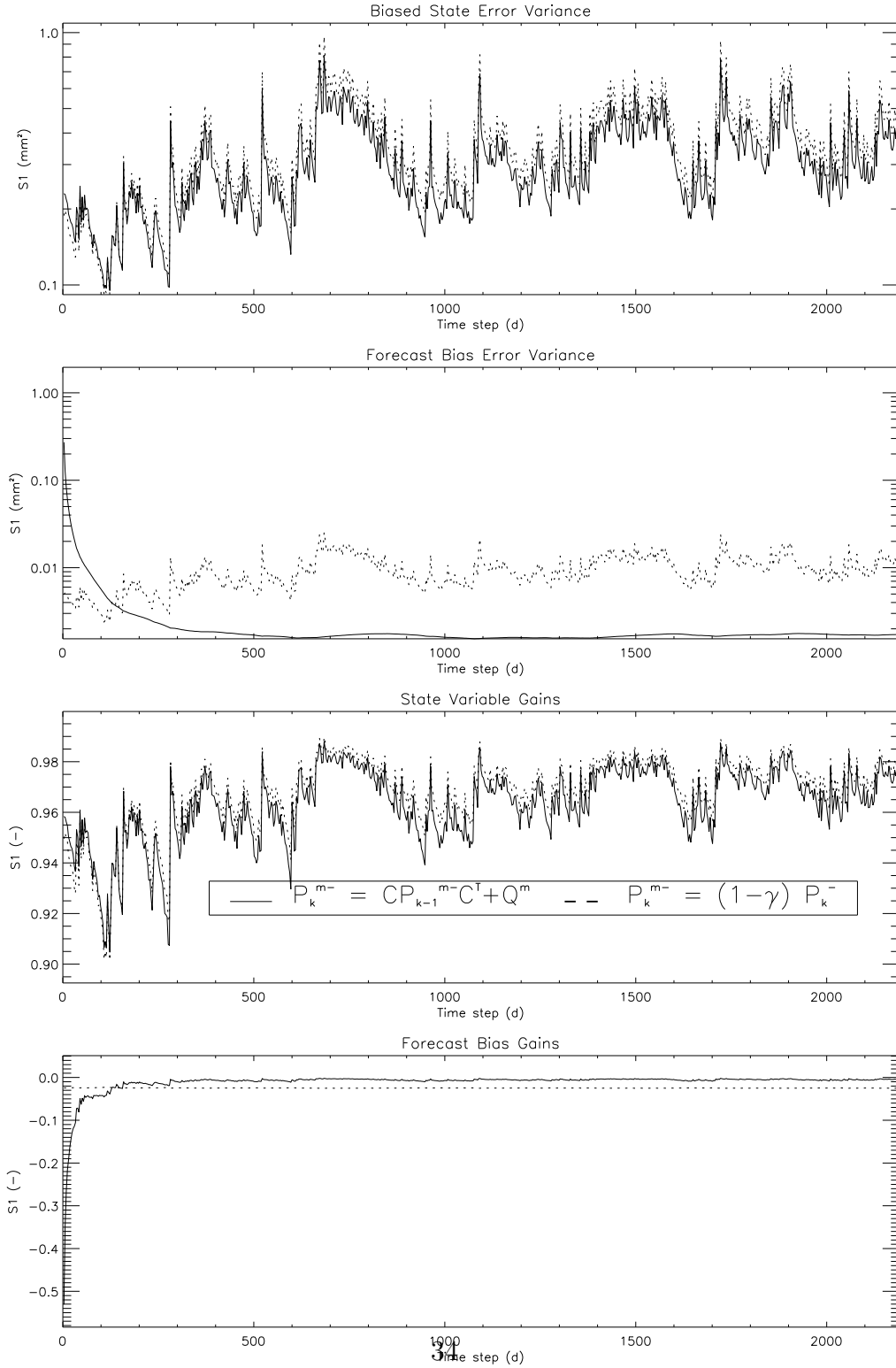


Figure 6: The a posteriori biased state and forecast bias error covariances, the gains for the state variables, and the gains for the forecast biases, during the validation period. The solid line represents the explicit calculation, the thick dashed line the simplified calculation.

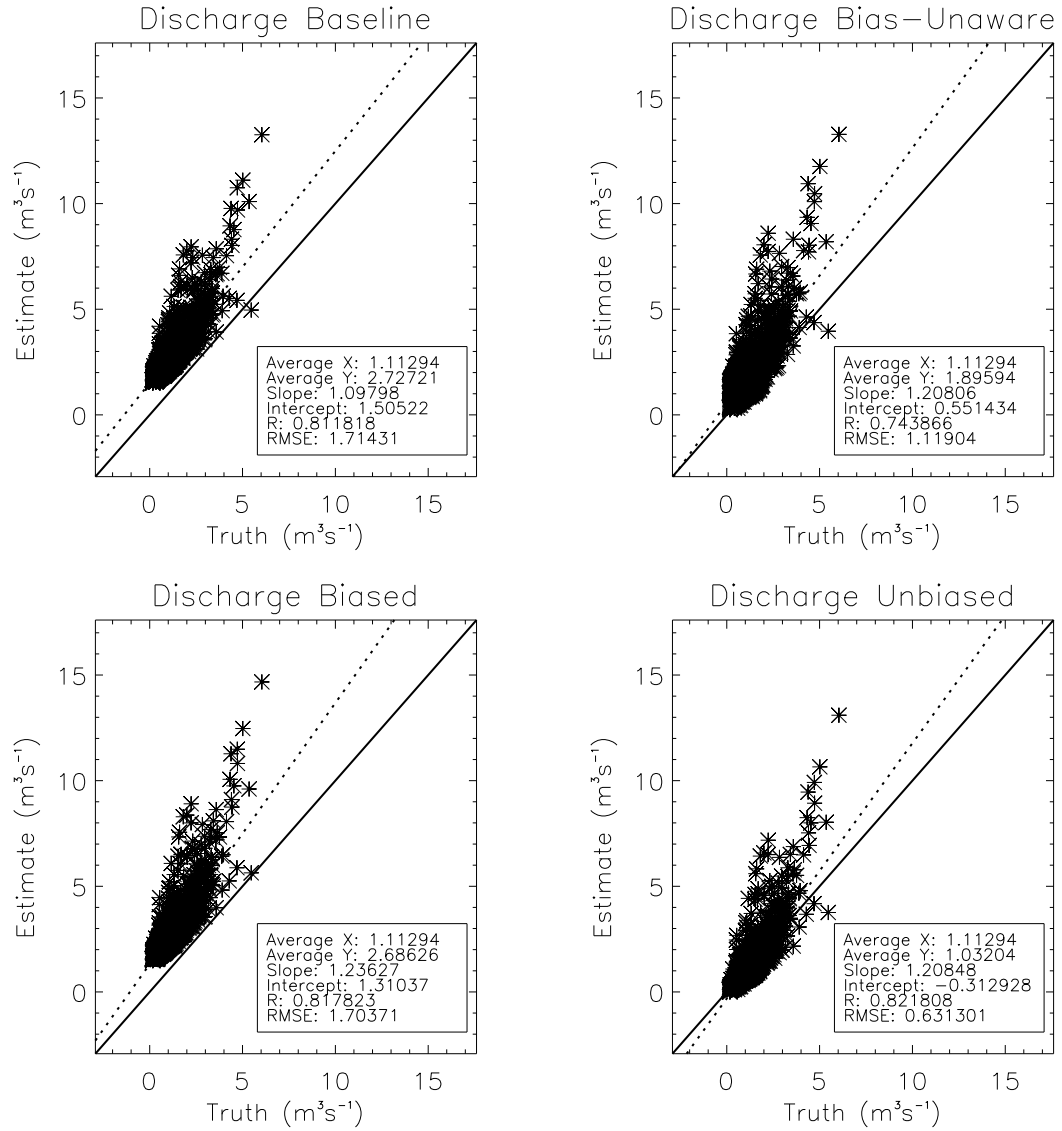


Figure 7: The modeled discharge for the persistent forecast bias model with explicit propagation of the forecast bias error covariance. Data are shown for the 6-year validation period following a 3-year calibration period. The top left hand side panel shows the results of the baseline run and the top right hand side the results of the bias-unaware filter. The bottom panels show the results of the bias-aware filter, with the left hand side panel the biased and the right hand panel side the unbiased estimates.

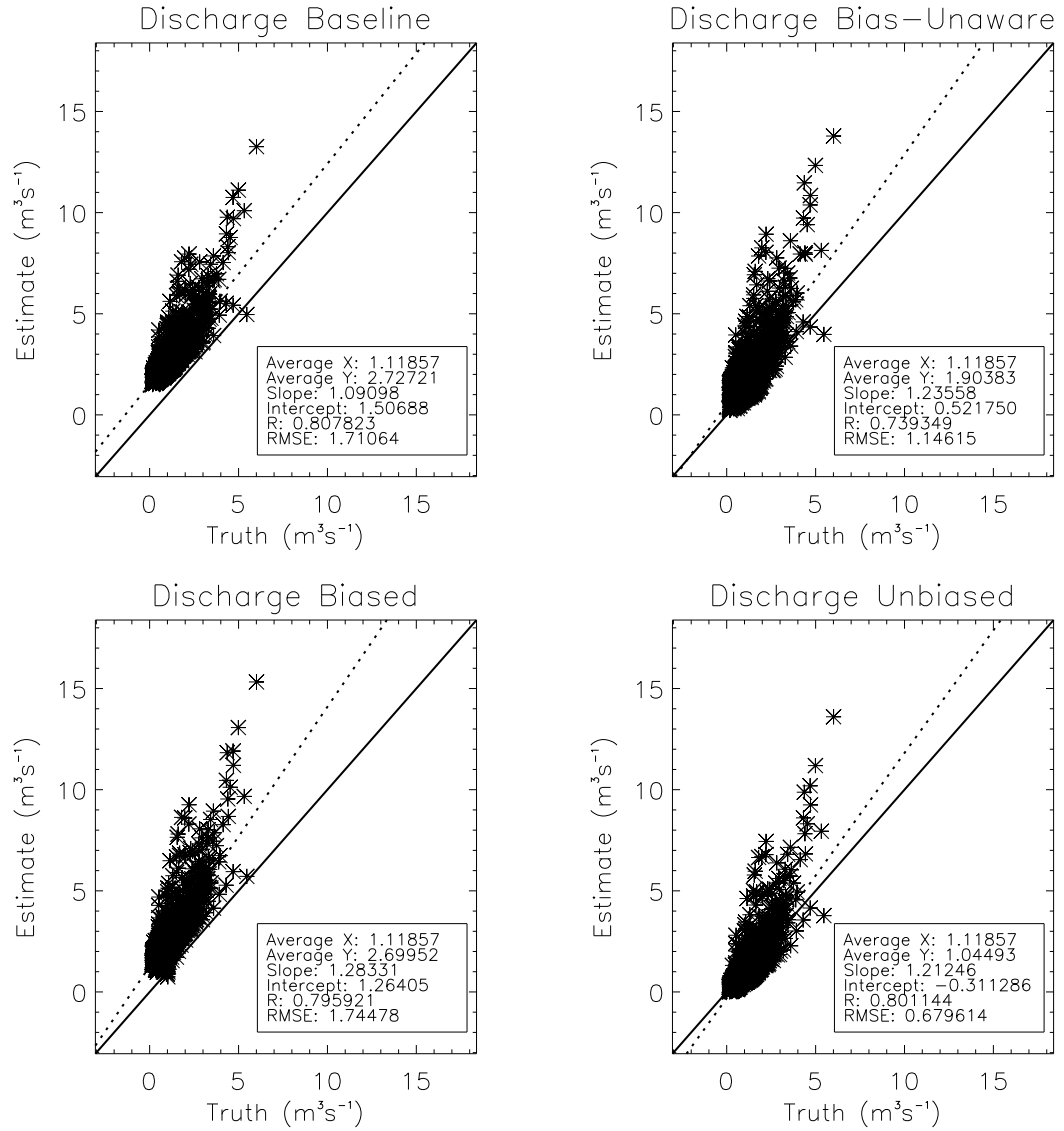


Figure 8: The modeled discharge for the persistent forecast bias model with simplified calculation of the forecast bias error covariance. Data are shown for the 6-year validation period following a 3-year calibration period. The top left hand side panel shows the results of the baseline run and the top right hand side the results of the bias-unaware filter. The bottom panels show the results of the bias-aware filter, with the left hand side panel the biased and the right hand side panel the unbiased estimates.

Appendix A. Propagation of the Forecast Bias Error Covariances

The error covariance of the forecast bias can be calculated as follows:

$$\mathbf{P}_k^{m-} = \mathbb{E} \left[\left(\mathbf{b}_k^m - \hat{\mathbf{b}}_k^{m-} \right) \left(\mathbf{b}_k^m - \hat{\mathbf{b}}_k^{m-} \right)^T \right] \quad (\text{A.1})$$

We substitute the propagation equation:

$$\mathbf{P}_k^{m-} = \mathbb{E} \left[\begin{array}{c} \left(\mathbf{C}_{k-1} \mathbf{b}_{k-1}^m + \mathbf{d}_{k-1} + \mathbf{w}_{k-1}^m - \mathbf{C}_{k-1} \hat{\mathbf{b}}_{k-1}^{m+} - \mathbf{d}_{k-1} \right) \cdot \\ \left(\mathbf{C}_{k-1} \mathbf{b}_{k-1}^m + \mathbf{d}_{k-1} + \mathbf{w}_{k-1}^m - \mathbf{C}_{k-1} \hat{\mathbf{b}}_{k-1}^{m+} - \mathbf{d}_{k-1} \right)^T \end{array} \right] \quad (\text{A.2})$$

This can be simplified as:

$$\mathbf{P}_k^{m-} = \mathbb{E} \left[\left(\mathbf{C}_{k-1} \mathbf{b}_{k-1}^m - \mathbf{C}_{k-1} \hat{\mathbf{b}}_{k-1}^{m+} + \mathbf{w}_{k-1}^m \right) \left(\mathbf{C}_{k-1} \mathbf{b}_{k-1}^m - \mathbf{C}_{k-1} \hat{\mathbf{b}}_{k-1}^{m+} + \mathbf{w}_{k-1}^m \right)^T \right] \quad (\text{A.3})$$

Since the bias and the noise vectors are uncorrelated, this reduces to:

$$\mathbf{P}_k^{m-} = \mathbf{C}_{k-1} \mathbf{P}_{k-1}^{m+} \mathbf{C}_{k-1}^T + \mathbf{Q}_{k-1}^m \quad (\text{A.4})$$